



OPEN Comparative genomic insights into multidrug resistance in classical and hypervirulent *K. pneumoniae* clinical isolates

Azra¹, Ihsan Ullah¹, Woranich Hinthong², Jody E. Phelan², Susana Campino², Ihtisham Ul Haq³, Zia Ud Din⁴, Sajjad Ahmad¹, Gulab Fatima Rani¹, Aneeqa Naz⁵, Otavio Cabral-Marques^{6,7,8,9}, Muhammad Zubair Bhatti¹⁰, Taj Ali Khan^{1,11}✉ & Taane G. Clark^{2,12}✉

Klebsiella pneumoniae is a highly adaptable opportunistic pathogen responsible for various infections, particularly in immunocompromised individuals. The recent genetic evolution of *K. pneumoniae* has led to the emergence of strains exhibiting both hypervirulence and multidrug resistance (MDR). In this study, we compared the genomic characteristics of MDR in classical and hypervirulent *K. pneumoniae* clinical isolates. Sixty-four isolates were collected from two healthcare institutions in Khyber Pakhtunkhwa, Pakistan, and identified using standard microbiological techniques. Whole-genome sequencing (WGS) was performed on 30 selected isolates using an Illumina platform. The sequencing data were analysed with the Kleborate tool. WGS analysis identified sequence types ST147 (n = 7, 23%) and ST2629 (n = 5, 16.6%) as the most prevalent lineages. Additionally, 13 other ST were detected, including ST147-1LV (n = 1, 3.3%) and ST859 (n = 1, 3.3%), which exhibited hypervirulent MDR traits. Among K-loci-associated virulence determinants, KL64 was the most predominant (n = 7, 23%), while the O-serotype O1/O2v1 was found in 22 isolates. A diverse range of antimicrobial resistance (AMR) determinants was observed across isolates. Plasmid-mediated quinolone resistance (PMQR) genes (*qnrS1*, *qnrB1*, *qnrB4*, *qepA2*) and aminoglycoside-modifying enzymes (*aac(3)-IIa*, *aac(6')-Ib-cr*, *strA*, *strB*) were detected in both classical and hypervirulent strains. β -lactamase genes included *bla*-SHV variants (*bla*-SHV-1, *bla*-SHV-11, *bla*-SHV-25, *bla*-SHV-187), *bla*-CTX-M-15, *bla*-VEB-5, *bla*-OXA-1, *bla*-OXA-48, and *bla*-OXA-181. Additional resistance genes conferred resistance to macrolides (*ermB*, *mphB*), phenicols (*catB4*, *catA1*, *floR*), sulfonamides (*sul1*, *sul2*), tetracyclines (*tetA*, *tetB*, *tetD*), and trimethoprim (*dfrA1*, *dfrA12*, *dfrA14*). Acquired AmpC and other β -lactamases (*DHA-1*, *TEM-1D*, *CMH-1*, *CMY-6*, *LAP-2*, *AmpC1*) were also present. The most prevalent plasmid replicons were Col(pHAD28), IncFIB(K), and IncR, with the hypervirulent ST147-1LV isolate carrying the highest number. These findings underscore the significant public health threat posed by hypervirulent and MDR *K. pneumoniae* strains, highlighting the extensive burden of virulence and resistance genes that complicate treatment strategies.

Keywords *K. pneumoniae*, hypervirulence, Resistance genes, Hybrid hvKp-MDR, Antimicrobial sequencing, WGS

¹Institute of Pathology and Diagnostic Medicine, IP&DM, Khyber Medical University Peshawar, Peshawar, Pakistan. ²Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK. ³Post-Graduate Programme in Innovation in Technology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais 30150-240, Brazil. ⁴Department of Pharmacy, Kohat University of Science and Technology, Kohat, Pakistan. ⁵Department of Pathology, Multan Medical and Dental College, Multan, Pakistan. ⁶Department of Immunology, Institute of Biomedical Sciences, University of São Paulo, São Paulo 05508-220, Brazil. ⁷Department of Medicine, Division of Molecular Medicine, Laboratory of Medical Investigation 29, University of São Paulo School of Medicine, São Paulo, Brazil. ⁸Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, Brazil. ⁹D'Or Institute for Research and Education, São Paulo, Brazil. ¹⁰Lady Reading Hospital, Peshawar, Pakistan. ¹¹Public Health Reference Laboratory, Khyber Medical University, Peshawar, Pakistan. ¹²Faculty of Epidemiology and Population Health, London, School of Hygiene and Tropical Medicine, London, UK. ✉email: Tajalikhan.ibms@kmu.edu.pk; taane.clark@lshtm.ac.uk

Klebsiella pneumoniae is a highly pathogenic Gram-negative bacterium known for its rapid mutation rate, which facilitates the emergence of hypervirulent and multidrug-resistant (MDR) strains^{1,2}. This relentless evolution has rendered many antimicrobials ineffective against *K. pneumoniae*¹. The pathogen is categorised into two distinct pathotypes: hypervirulent *K. pneumoniae* (hvKp) and classical *K. pneumoniae* (cKp)². These pathotypes differ molecularly, geographically, and clinically. The World Health Organization (WHO) has identified cKp isolates as a “critical concern” because they frequently cause nosocomial infections in immunocompromised individuals or those with underlying illnesses, and they have progressively gained antibiotic resistance determinants, primarily carbapenemases^{3–5}. However, hvKp has become a virulent pathogen in recent decades that can infect healthy people with invasive community-acquired illnesses. Rapid metastatic spread and the development of pyogenic tissue abscesses are characteristics of these infections⁵.

cKp is predominantly associated with nosocomial infections, and hvKp is responsible for community-acquired infections². In *K. pneumoniae*, both classical and hvKp strains exhibit significant AMR and virulence, complicating infection severity². Classical strains are often linked to nosocomial infections and display multidrug resistance through mechanisms like extended-spectrum beta-lactamases (ESBLs)^{2,6}. Hypervirulent strains possess aggressive pathogenic traits and carry antibiotic resistance profiles^{2,7,8}. This combination of heightened virulence and robust resistance in hypervirulent strains makes them particularly challenging to treat^{7,8}.

Classical and hypervirulent *K. pneumoniae* strains possess key virulence factors, such as capsular polysaccharides (CPS) and lipopolysaccharides (LPS), encoded by K-loci and O-loci, respectively. To date, 134 K-types and 11 O-antigens have been identified^{9,10}. Both CPS and LPS can suppress early inflammatory responses, defend the bacteria against phagocytosis, opsonisation, and antimicrobial peptides, and stimulate the host's innate immune response¹¹. The heightened immunogenicity and prominent surface exposure of CPS and LPS in *K. pneumoniae* make them candidates for vaccine development¹². CPS types KL1 and KL2 are linked to increased virulence, while KL47 and KL64 are associated with both hypervirulence and carbapenem resistance¹³. Diversity in capsular types has been identified, with most types associated with specific sequence types (STs); for example, ST11 with K27 or K64, ST101 with K17, ST340 with KL151, ST15 with K24, and ST17 with KL112¹⁴. The structural diversity of the O-antigen enhances immune evasion¹⁵. The plasmids often harbour genes essential for virulence and antibiotic resistance^{16,17}. Genomic surveillance is increasingly used to monitor hypervirulence and carbapenemase genes in *K. pneumoniae*, particularly in high-risk clonal complexes like CC258 (ST11, ST258, ST512, and ST437)^{1,18}. The emergence of hypervirulent and carbapenem-resistant strains (CR-hvKp) has become a significant global health threat^{19–21}. CR-hvKp infections are challenging to treat with available medicines and warrant more investigation^{19–21}.

Whole Genome Sequencing (WGS) has emerged as a critical tool for identifying and monitoring the genetic determinants of bacterial virulence and resistance, and enables a deeper understanding of the mechanisms underlying the pathogen's virulence, resistance, and adaptability^{2,4,6,21,22}. Carbapenem resistance in *K. pneumoniae* mainly results from carbapenemase production (e.g., KPC, NDM, OXA-48), often combined with porin loss and efflux pump activation, which together reduce antibiotic efficacy and lead to high-level resistance^{23,24}. This research aims to conduct a genomic analysis of classical and hypervirulent *K. pneumoniae* strains from Pakistan to elucidate the diversity of genes associated with virulence, antibiotic resistance, and plasmids across different sample sources.

Materials and methods

Sample collection and pathogen isolation

The study included 500 samples sourced from urine, pus, wounds, blood, and other body fluids collected from patients of Mardan Medical Complex (Mardan) and Ayub Medical Complex (Abbottabad), Khyber Pakhtunkhwa (KPK), Pakistan, to isolate *K. pneumoniae*. Informed consent was obtained from all participants, and ethical approval was granted. The samples were aseptically inoculated onto sterile Blood and MacConkey agar (Oxoid, UK) medium and incubated at 37 °C for 24 h. *K. pneumoniae* were identified by biochemical profiling, and strain confirmation was determined using the API 20E kit (bioMérieux SA, Lyon, France).

Genomic DNA isolation and whole genome sequencing

K. pneumoniae preserved isolates were meticulously inoculated onto Mueller–Hinton Agar plates and incubated for 24 h at 37 °C. DNA extraction was performed using the DNeasy UltraClean microbial kit (Qiagen, Germantown,

MD, USA). DNA quantification was performed using the Qubit dsDNA fluorometer (ThermoFisher Scientific, Waltham, MA, USA). Based on the phenotypic susceptibility profiles, thirty samples with high-quality DNA underwent WGS at The Applied Genomics Centre, London School of Hygiene and Tropical Medicine. Before library preparation and sequencing, DNA quantity and quality were re-evaluated. Libraries of bacterial genomes were prepared using the Illumina library preparation kit (San Diego, CA, USA), following the manufacturer's protocol. Sequencing was conducted using the Illumina MiSeq platform, generating paired-end short reads with an average length of 150 bp and approximately 50-fold coverage.

Data processing and quality control

The FastQC (v0.11.9) tool was employed to conduct comprehensive pre-processing of FastQ files²⁵, and Trimmomatic (v0.36) software was used to remove adapter sequences²⁶. Data-quality reports were generated pre- and post-trimming and then consolidated into a single report using MultiQC. Before genome assembly, paired-end reads underwent taxonomic classification with Kraken2 software²⁷, employing default parameters to determine species identity and assess potential contamination. Kraken2 outputs were visualised using a Krona chart, providing a graphical overview for contamination assessment. Processed paired-end reads were then assembled de novo using the SPAdes genome assembler²⁸ within the Shovill pipeline (<https://github.com/tseemann/shovill>; accessed on 20 May 2024) with default settings. Assembly statistics, such as N50 values and the number of contigs, were calculated with QUAST using default settings²⁹.

Genomic profiling

The de novo assembled sequences served as the basis for strain characterisation. Species verification was cross-referenced using the online Type Strain Genome Server (TYGS; <https://tygs.dsmz.de>; accessed on 15 June 2024). To investigate antibiotic resistance genes and the presence of mobile genetic elements, the web tools ResFinder (v4.1) and PlasmidFinder (v1.0 and v2.1) were employed^{30,31}. Serotyping, cgMLST and cSNP-phylogenetic analyses were conducted using the online Pathogenwatch platform (<https://pathogen.watch>; accessed on 10 June 2024). The MLST scheme from PubMLST was utilised to assign a sequence type (ST) to each strain. For the validation of AMR detection, the ResFinder database was applied for alignment-based analysis. Assembled genomes were aligned against the database using BLASTn, with only alignments exhibiting more than 90% identity and over 60% target coverage being retained. Genes related to virulence were detected using ABRicate (<https://github.com/tseemann/abricate>; accessed on 3 June 2024) using the Virulence Factor Database (VFDB)³². Along with species identification, we used Kleborate software to methodically evaluate the existence of virulence determinants and resistance, as well as STs and virulence scores³³.

Genome annotation and phylogenetics

Genome annotation was conducted using Prokka software, which delineated features such as coding sequences (CDS) and ribosomal and transfer RNA genes. Various standard output files were generated for subsequent analyses. The Roary pipeline utilised Prokka's GFF-generated files to construct pan-genomes and perform core-genome-based phylogenetic analysis³⁴. The core genome was defined based on the criteria of 95% identity for protein matches and the presence in 99% of strains. A multi-fasta alignment of all isolates was performed to construct a phylogenetic tree using IQ-TREE with default parameters. Subsequently, a core genome-based phylogeny was constructed to facilitate the clustering of all isolates and visualised in the iTOL³⁵. Variant calling and filtering were performed using the Snippy pipeline (v4.6.0), leading to a SNP-based phylogeny.

Results

Genomic analysis and assembly statistics of *K. pneumoniae* strains

A total of sixty-four *K. pneumoniae*-positive samples were analysed. DNA extraction was performed, and thirty samples based on phenotypic profile and high DNA quality were selected for further processing (Suppl. Table 1, Suppl. Table 2). The WGS of *K. pneumoniae* were performed using the Illumina MiSeq platform. Assembly statistics of all the genomes are detailed (Suppl. Table 3). The genomes exhibited between 5,192 and 11,192 protein-coding sequences (CDS), 3 to 23 ribosomal RNA genes (rRNAs), and 72 to 149 transfer RNA genes (tRNAs). Their sizes ranged from 5,332,556 to 11,152,258 bp, with GC contents varying from 53.99% to 57.44%. Assembly quality metrics showed N50 values between 21,500 and 45,000 bp, with corresponding L50 values of 25 to 60, indicating reliable assembly continuity. Sequencing depth ranged from 50- to 150-fold, ensuring high-quality base coverage, and annotation revealed plasmid sequences and mobile genetic elements in several isolates. (Suppl. Table 3).

Identification of high-risk sequence types and novel k-serotypes

The isolates exhibited significant genetic diversity, with 15 distinct STs identified. Among these, high-risk STs such as ST147, ST2629, and ST4 were noted, alongside other types including ST147-1LV, ST1310, and ST37 (Fig. 1, and Suppl. Table 4). The study identified 14 K-serotype groups that characterise the *K. pneumoniae* population, primarily associated with the K loci. Observed capsule types included KL64 (n = 7), KL16 (n = 5), and KL114 (n = 1).

Capsule types KL17, 19, 51, 64, 114, 127, and 151 are associated with hypervirulence, whereas KL2 and 31 are indicative of high-risk MDR clones, and KL13, 16, 48, and 62 are characteristic of classical *K. pneumoniae*. Several isolates belonged to K serotypes where the serological capsule groups could not be precisely identified. Among the eight O-serotype groups identified, the predominant serotypes were O1/O2v1 (n = 22) and O1/O2v2 (n = 3). The remaining serotypes O3b (n = 2), OL13 (n = 2), O5 (n = 3), and OL104 (n = 1) were found at lower frequencies within the *K. pneumoniae* isolate population (Fig. 2).

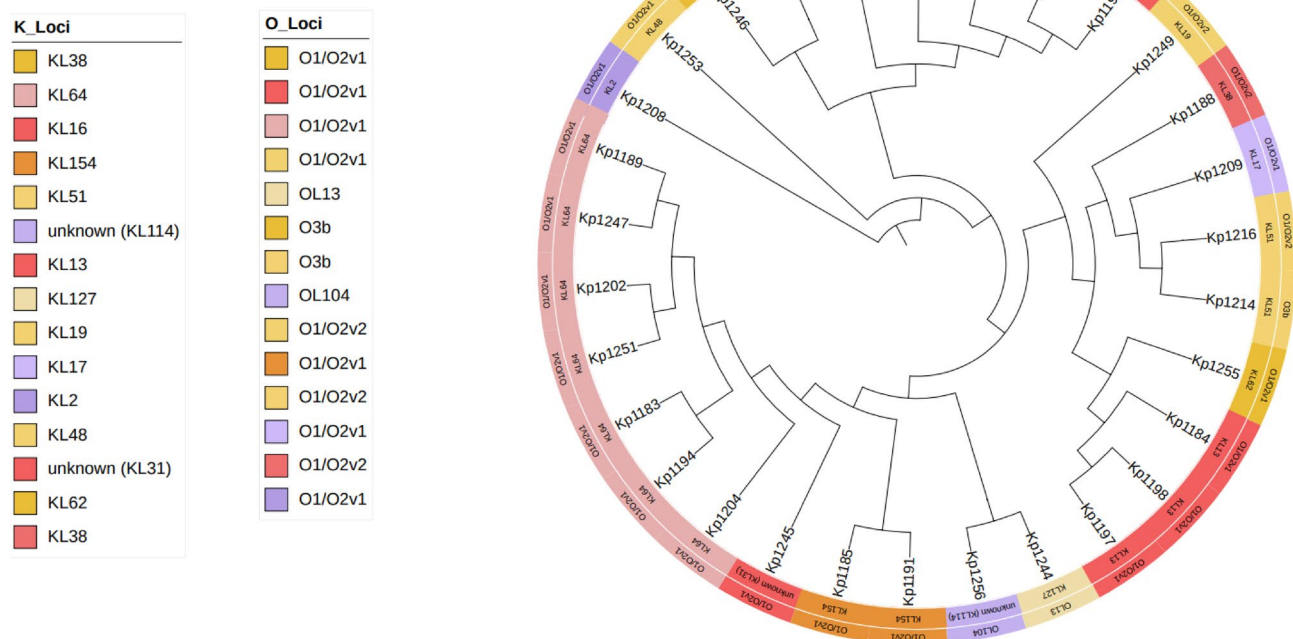


Fig. 1. K loci and O loci distribution in the 30 isolates.

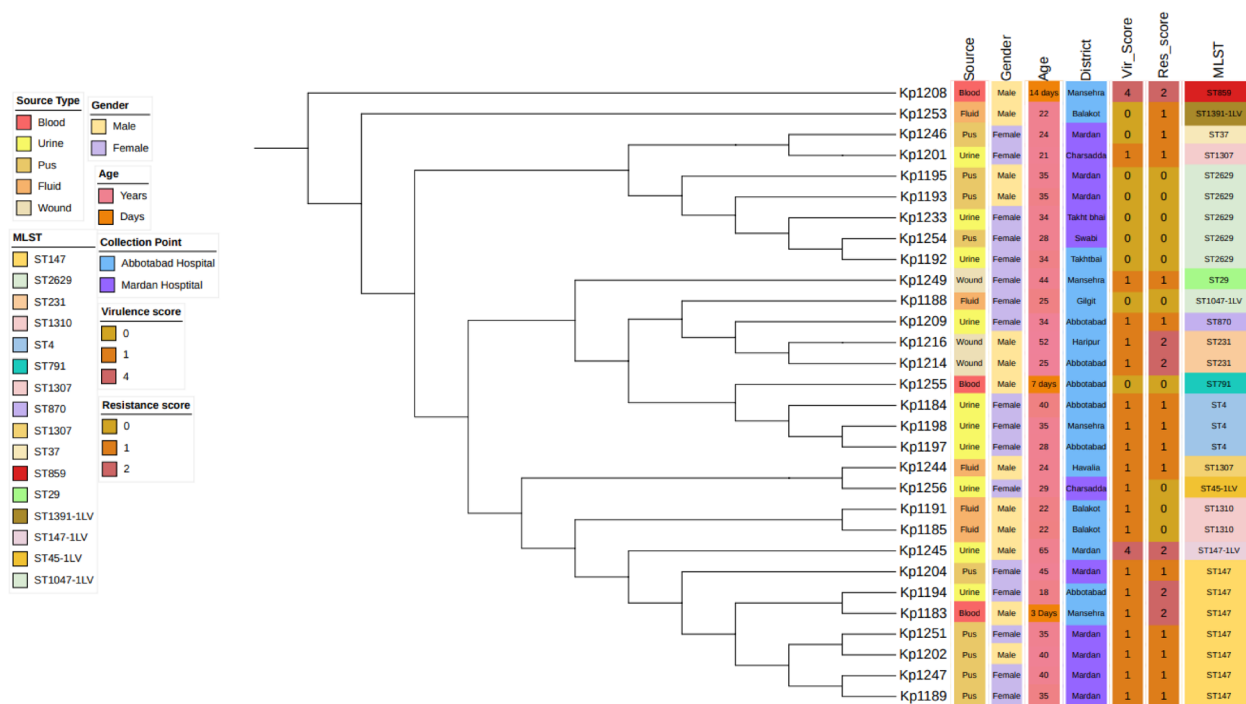


Fig. 2. Phylogenetic tree based on core genes from 30 *K. pneumoniae* isolates, annotated with isolation source, gender, age, district, MLST, virulence score, resistance score, K loci, and O loci, generated by iTOL³⁵.

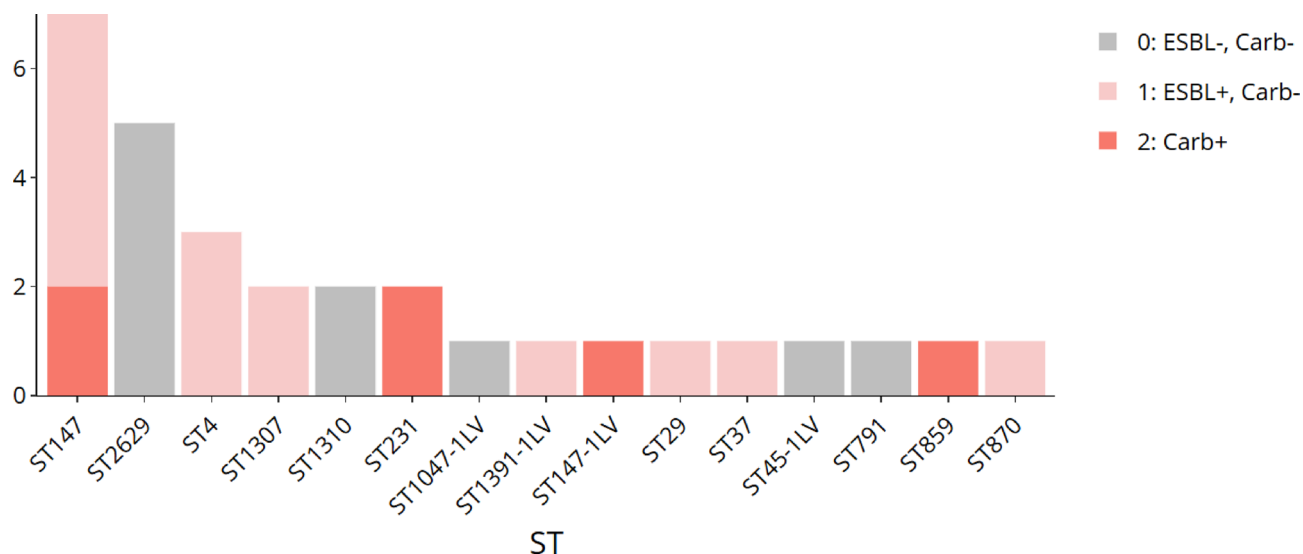


Fig. 3. Isolate STs and resistance profiles as determined by Kleborate software³³.

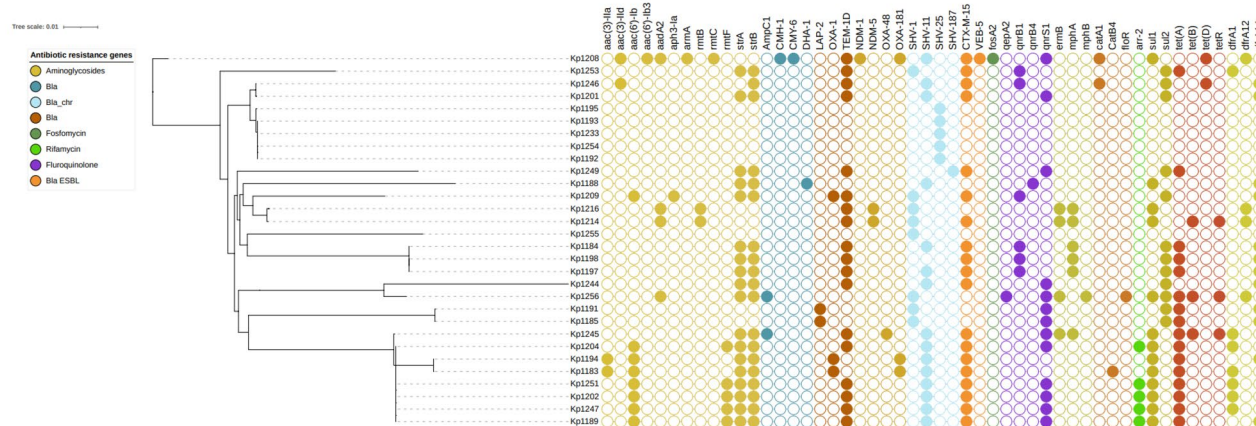


Fig. 4. Core gene-based phylogenetic tree of 30 *K. pneumoniae* isolates with their corresponding resistance gene profiles. Filled circles indicate gene presence, while empty circles denote absence, as visualised in iTOL³⁵.

Identification of different AMR patterns

Antibiotic resistance genes were identified across nearly all isolates. Every strain exhibited resistance to SHV beta-lactamase; however, some did not exhibit changes at position 35Q. A diverse array of carbapenemase genes were detected, including *bla*-NDM-1 ($n=1$, 3.3%), *bla*-NDM-5 ($n=2$, 6.6%), along with other lactamases like *bla*-OXA-1 ($n=3$, 9.9%), and *bla*-CTX-M-15 (63.3%). The ST859 harboured most of these carbapenemase genes, with a significant presence also in ST147-1LV, ST870, and ST1047. The high-risk clone ST859 harboured a broad repertoire of acquired and chromosomal resistance genes, including *CMH-1*, *CMY-6*, and *TEM-1D.v1*, and *SHV-11* (Suppl. Table 5). The ten strains lacking ESBLs comprised five of ST2629, one of ST1047-1LV, two of ST1310, one of ST791, and one of ST45-1LV (Fig. 3). Strains devoid of *bla*-TEM-1D included two of ST147, two of ST1310, five of ST2629, one of ST791, and one ST147-1LV (Fig. 3).

Nine strains harboured more than three aminoglycoside resistance genes. Interestingly, twenty isolates exhibited phenotypic susceptibility to aminoglycosides despite the presence of corresponding resistance genes (Suppl. Table 6). The ST859 strain uniquely carried the *armA* gene. Among the 30 isolates, only one harboured the aminoglycoside phosphotransferase gene *aph3*. The most prevalent aminoglycoside resistance determinant was the plasmid-encoded streptomycin resistance gene B, *strB* ($n=19$, 63%). Other identified resistance genes included acetyltransferase *aac*³-IIa *aac*(6')-Ib, *strA.v1*, and *aadA2*. (Fig. 4, Suppl. Table 5). Eighteen strains possessed a single gene, while two exhibited two genes linked to fluoroquinolone resistance. The most prevalent gene, *flq qnrS1*, was detected in 14 strains, including two in ST1310, two in ST1047-1LV, and six strains harboured the *flq qnrB1* gene. Twelve strains possessed the *Slusul1* gene, whereas 13 strains carried the *Slusul2* gene (Fig. 4, Suppl. Table 5). Nineteen strains were found to harbour genes conferring resistance to tetracyclines. Among these, 16 strains possessed the *tet*(A) gene, two strains carried *tet*(D), and one strain contained *tet*(B).

v2. Each resistant strain was associated with a single resistance gene. The *tet(D)* gene was present in ST859 and ST37, while ST231 harboured *tet(B)* (Fig. 4, Suppl. Table 5). The *drfA12* gene was found in four strains: one ST859 strain, two ST231 strains, and one ST45-1LV strain. Strains ST2629, ST29, and ST791 were negative for trimethoprim resistance genes (Fig. 4, Suppl. Table 5). *Oxa-48* and *Oxa-181* were detected in different STs.

Demonstration of rare resistance by some STs

Among the analysed strains, seven exhibited macrolide resistance, with four carrying the *ermB* gene and three harbouring the *mphA* gene, all within the ST4 lineage. Notably, only five of the ST147 strains possessed the *arr-2* gene, indicating rifampicin resistance, while the remaining strains were free from rifampicin resistance. The *fosA2* gene, associated with fosfomycin resistance, was found exclusively in ST859. Comprehensive resistome and phenotypic antibiogram data for all strains are presented (Fig. 4, Suppl. Table 5, Suppl. Table 6).

Comprehensive analysis of enzyme alterations and genetic mutations

Mutations were observed in the Sulphydryl Variable (SHV-11, SHV-25) enzyme. Mutations in *OmpK35* were detected in 7% of ST147 strains and 30% of ST231 strains. Concurrently, a mutation in *OmpK36* resulted in aspartic acid (D) replacing glycine (G) in all ST231 strains (Suppl. Table 5). However, the ST147-1LV strain exhibited multiple mutations in *gyrA*, such as the replacement of serine (S) with isoleucine (I) and leucine (L) at position 83 and aspartic acid (D) with asparagine (N) at position 87 (Suppl. Table 5). A truncated, unaltered *CTX-M-15* gene was found in ST231, while the *strA* (streptomycin resistance gene) was identified in ST37 (Suppl. Table 5).

Prevalence of virulence-associated factors

Six strains tested positive in the string test, indicating the presence of multiple hypervirulent strains. Yersiniabactin emerged as the most prevalent virulence factor, found in 22 isolates. The most common variant was *ybt16*, associated with *ICEKp12* and identified in 8 isolates, particularly in ST147 strains (with one strain carrying a truncated gene), as well as in ST859. Three isolates harboured *ybt10* linked to *ICEKp4*, found in ST1310 strains and ST45-1LV. An incomplete *ybt9* associated with *ICEKp3* was present in ST147-1LV. Only two strains contained Aerobactin genes: ST859 carried *iuc1*, and ST147-1LV harboured *iuc5*. Neither colibactin nor salmochelin was detected in any isolates (Figs. 2 and 5, and Suppl. Table 4). Hyper-mucoviscosity, governed by the regulator of mucoid phenotype A (*rmpA*) and the transcriptional activator (*rmpA2*) genes, is closely associated with hypervirulence².

Virulence and resistance score

This study detected a maximum of 67 virulent genes across the isolates. Notably, the acquired virulence genes coding for colibactin, nutrition factors, salmochelin, and *rmpA* were absent. Yersiniabactin, an iron uptake locus (*ybtAEPQSTUX*), was identified in 21 (70%) isolates, while the regulatory gene *rmpA2* was present in only one isolate. Aerobactin, a dominant siderophore, was found in two (6.6%) isolates belonging to ST859 and ST147-1LV. Isolates with a resistance score of 0 belonged to ST1310, ST1047, ST2629, ST791, and ST45-1LV. The AMR score was calculated based on the presence of ESBL (1 score), carbapenemase without colistin resistance (2 scores), and carbapenemase plus colistin resistance (3 scores), with all other cases scoring zero (Figs. 2 and 5, Suppl. Table 4, Suppl. Table 5).

Prevalence and distribution of plasmid replicon families

Among the 28 replicon families identified by PlasmidFinder software, three were present in more than 45% of the isolates: IncFIB(K) in 57%, Col(pHAD28) in 50%, and IncR in 47%. Notably, ST147-1LV exhibited the highest number of replicons¹⁶, followed by ST231 with eleven and ST859 with nine replicons. On average, each

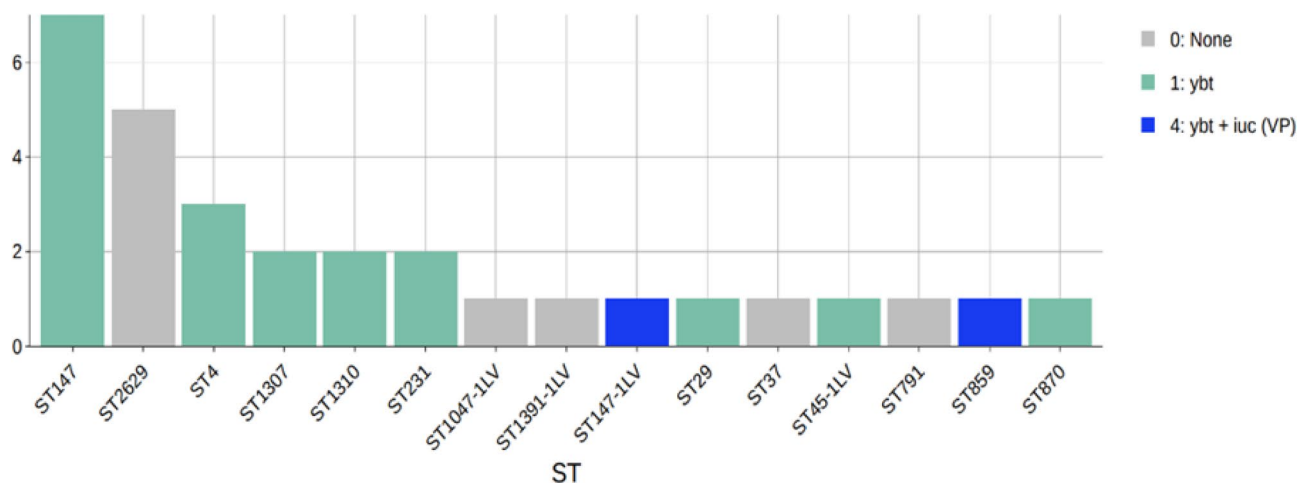


Fig. 5. Isolate STs and virulence scores as determined by Kleborate software³³.

isolate contained five replicons. Specifically, three ST2629 and one ST1307 isolates each harboured a single replicon (Fig. 6, Suppl. Table 6).

Phylogenetic analysis of *K. pneumoniae* isolates

The phylogenetic analysis of *K. pneumoniae* isolates (Fig. 7), revealed substantial genetic diversity. The core genome of the *Klebsiella* dataset comprises 3,935 genes. Four distinct clusters were identified from the analysed genomes. The ST859 and ST1391-1LV strains formed a unique cluster, distinguishing them from all other strains. Another cluster included ST37, ST1307, and ST2629. A third cluster comprised ST791, ST4, and ST29, along with isolates of ST231, ST870, and ST1047-1LV. The final cluster, the largest phylogroup, comprised six groups formed by the remaining 12 isolates, including sequence types ST147, ST147-1LV, ST1307, and ST45-1LV.

Discussion

Globally, AMR is regarded as a “silent pandemic” and a public health issue, especially in low- and middle-income countries (LMICs) like Pakistan³⁶. This study elucidates the genetic characteristics of hvKp and cKp isolates recovered from clinical settings. The current study examined *K. pneumoniae* sourced from blood, pus, and urine, comprising several STs. *K. pneumoniae* was most frequently detected in urine, which aligns with the observations reported by others³⁷. There has been a significant increase in the prevalence of MDR and Extensively Drug-Resistant (XDR) *K. pneumoniae*, estimated to be over 40% in Pakistan³⁸. Notably, carbapenem-resistant *K. pneumoniae* (CRKP) has become more prevalent in Pakistan³⁹. Our research supports earlier findings that suggested numerous AMR genes linked to resistance to different medication classes were present in *K. pneumoniae*^{6,22,40}. Clinical isolates of *K. pneumoniae* may have more genetic and phenotypic diversity due to the accumulation of various antibiotic resistance genes brought about by horizontal gene transfer^{1,16}. The high-risk clone *K. pneumoniae* ST147, known for its association with MDR and pan-drug resistance in chronic infections, was detected in our dataset⁴¹. In Tuscany, Italy, an epidemic caused by a hypervirulent *K. pneumoniae* ST147 strain carrying a chimeric plasmid harbouring both virulence and resistance genes⁴² was reported, consistent with our findings.

Considering the *bla*-SHV gene has allelic variations in the core genome, all of the isolates in our research were ampicillin-resistant. The two most common SHV variants within them were *bla*-SHV-1 and *bla*-SHV-11⁴³. Several studies conducted in China have found this variation⁴⁴. Recent investigations have shown the variation *bla*-SHV-187 among MDR *Klebsiella* ST29 isolates^{45–47}.

Likewise, our analysis revealed the presence of genes such as *bla*-CTX-M-15, *bla*-CMH-1, and *bla*-OXA, linked to beta-lactam resistance. Furthermore, class B (*bla*-NDM gene NDM-1) and class D (*bla*-OXA, including OXA-48 and OXA-181) carbapenemase variants were detected; these variants have been described in isolates from Pakistan⁴⁸. The most prevalent PMQR genes found in our isolates were *qnrS1* and *qepA2*, which is an observation in line with earlier studies (e.g.,⁴⁹). We discovered that quinolone-resistant isolated strains had mutations in both *gyrA* and *parC* genes, similar to those identified in another study⁵⁰. In the *gyrA* gene, among the most common mutations were I83Serine, L83Serine, and D87Asparagine, while the *parC* gene had I80S. The most prevalent aminoglycoside-resistant genes identified in our investigation included *aac*(6′)-Ib-cr, and *aac*³-IId, *armA*, *strA*, *strB*, and *rmtB*, which contributed to high-level aminoglycoside resistance^{51,52}. The *ermB* and *mphA* genes were found to be associated with resistance to macrolides, lincosamides, and streptogramin B (MLS_B) antibiotics⁵³.

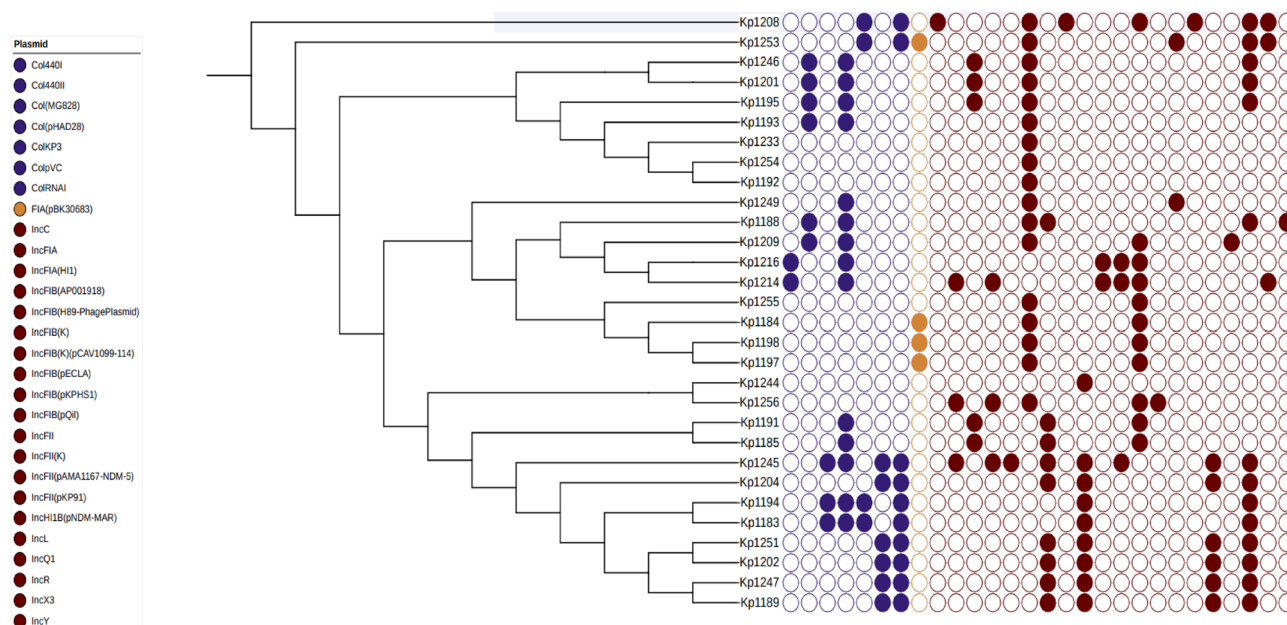


Fig. 6. Core genome-based Phylogenetic tree of 30 *K. pneumoniae* isolates with their plasmid replicons. Empty shapes show the absence of a plasmid; filled shapes show the presence of a plasmid, generated by iTOL³⁵.

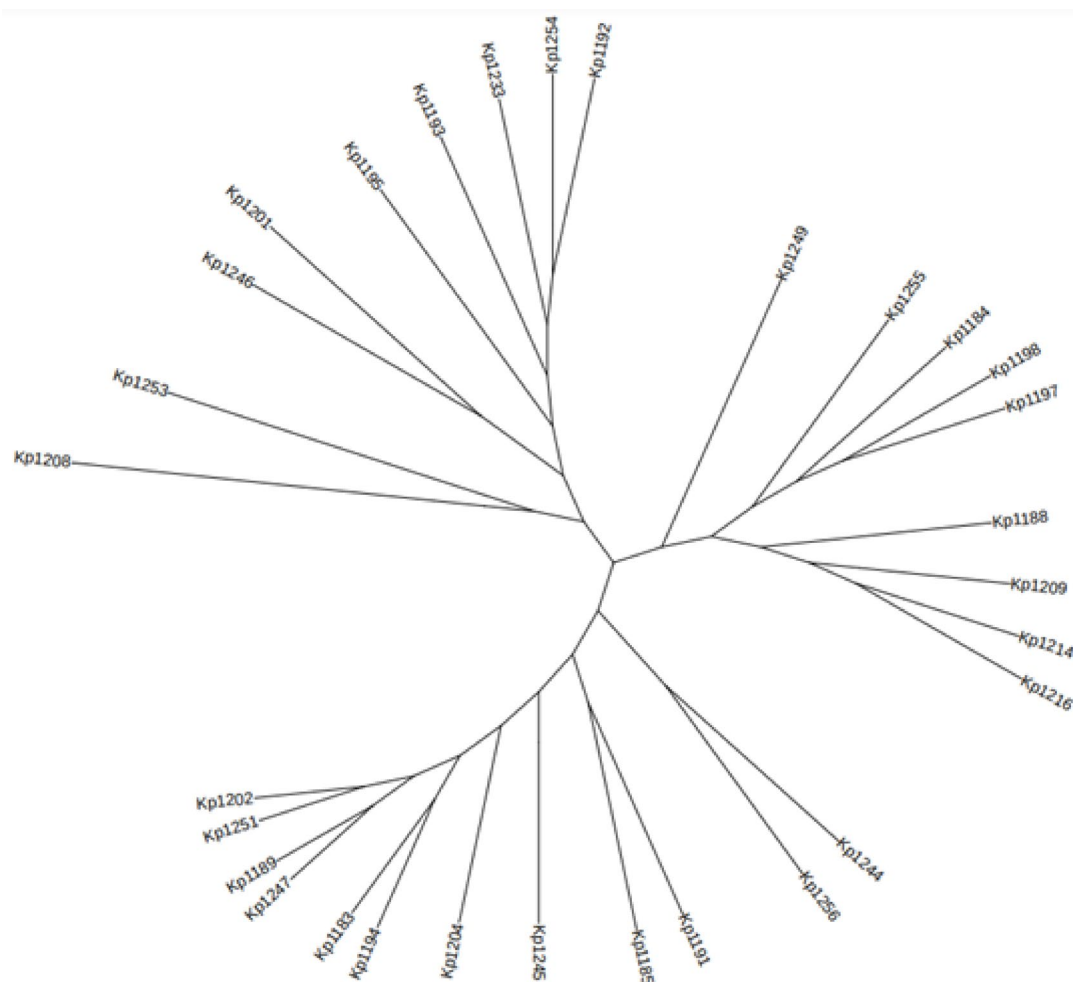


Fig. 7. Core genome SNP-based phylogenetic tree of *K. pneumoniae* isolates, generated with FigTree.

The biocide-resistant gene *qacE-1*, typically present in gram-negative microbes, is widely prevalent in our study. This gene is associated with resistance to sulfonamides, trimethoprim, and aminoglycosides via a multidrug efflux mechanism⁵⁴.

The presence of ESBL genes in comparable genetic environments in isolates of our study and other studies suggests that they promote antibiotic resistance⁵⁵. The insertion sequence *ISEc9* and specific plasmids, particularly those belonging to the IncF family, have been implicated in mediating gene exchange mechanisms that promote the dissemination of ESBL genes⁵⁶. The IncF plasmid, which plays a role in horizontal gene transfer, was identified in our strains⁵⁶. Only the ST859 strain carried an IncC plasmid replicon containing genes conferring resistance to trimethoprim, AmpC β -lactamase, sulfonamides, aminoglycosides, and chloramphenicol⁵⁷. In our study, the ST859 strain carried Col, FIA, and Inc-type plasmids (Inc(L, Q, R, X, Y)), along with the globally disseminated plasmid IncH1B(pNDM), which has previously been reported in Morocco, the United States, and the United Kingdom⁵⁷. The most common sequence type identified in our study was ST147, which exhibited the serotype combination KL64 and O2v1 across all seven strains. *K. pneumoniae* isolates from hospitalised patients have been reported to belong to ST147 in association with the KL64 capsular type⁵⁸. *K. pneumoniae* can be classified into 134 capsular types based on variations in polysaccharide composition, corresponding to different K and O antigens⁵⁹. It is well established that hypervirulent *K. pneumoniae* variants are associated with capsular type K54. This type is particularly linked to sequence type ST29, a hypervirulent strain commonly found within clonal group CG29⁶⁰.

Conclusion

We present a comprehensive WGS study of *K. pneumoniae* strains from clinical sources in Khyber Pakhtunkhwa, Pakistan, revealing substantial genetic diversity. Notably, three isolates - absent from the TYGS database - harboured key virulence genes and antimicrobial resistance determinants, indicating the presence of rare or underreported variants. The widespread occurrence of virulence and resistance genes highlights their potential contribution to poor therapeutic outcomes and elevated mortality rates. Furthermore, uncommon sequence types, K loci, and spontaneous mutations were detected, emphasising the need for ongoing epidemiological and molecular investigations to inform effective treatment strategies.

Data availability

The WGS data generated in this study have been deposited in the NCBI Sequence Read Archive under BioProject accession number PRJNA1148473. The isolates have been deposited in Pathogenwatch, available from https://pathogen.watch/collection/edexwe3w44o9-kleb_pak.

Received: 25 May 2025; Accepted: 31 October 2025

Published online: 06 December 2025

References

- Holt, K. E. et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc Natl Acad Sci USA* **112**(27), E3574–E3581 (2015).
- Spadar, A., Perdigão, J., Campino, S. & Clark, T. G. Genomic analysis of hypervirulent *Klebsiella pneumoniae* reveals potential genetic markers for differentiation from classical strains. *Sci Rep.* **12**(1), 13671 (2022).
- WHO Bacterial Priority Pathogens List. 2024: *bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance* (World Health Organization, 2024).
- Spadar, A., Perdigão, J., Campino, S. & Clark, T. G. Large-scale genomic analysis of global *Klebsiella pneumoniae* plasmids reveals multiple simultaneous clusters of carbapenem-resistant hypervirulent strains. *Genome Med.* **15**(1), 3 (2023).
- Choby, J. E., Howard-Anderson, J. & Weiss, D. S. Hypervirulent *Klebsiella pneumoniae*—clinical and molecular perspectives. *J. Intern. Med.* **287**(3), 283–300 (2020).
- Perdigão, J. et al. Whole-genome sequencing resolves a polyclonal outbreak by extended-spectrum beta-lactam and carbapenem-resistant *Klebsiella pneumoniae* in a Portuguese tertiary-care hospital. *Microb Genom.* **7**(6), 000349 (2019).
- Han X, Yao J, He J, Liu H, Jiang Y, Zhao D, et al. Clinical and laboratory insights into the threat of hypervirulent *Klebsiella pneumoniae*. *International Journal of Antimicrobial Agents.* 2024:107275.
- Beig, M. et al. Antibiotic resistance rates in hypervirulent *Klebsiella pneumoniae* strains: a systematic review and meta-analysis. *J Glob Antimicrob Resist.* **38**, 376–388 (2024).
- Wyres, K. L. et al. Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microb Genom.* **2**(12), e000102 (2016).
- Clarke, B. R. et al. Molecular basis for the structural diversity in serogroup O2-antigen polysaccharides in *Klebsiella pneumoniae*. *J. Biol. Chem.* **293**(13), 4666–4679 (2018).
- Flores-Valdez, M. et al. Whole genome sequencing of pediatric *Klebsiella pneumoniae* strains reveals important insights into their virulence-associated traits. *Front. Microbiol.* **12**, 711577 (2021).
- Choi, M. et al. The diversity of lipopolysaccharide (O) and capsular polysaccharide (K) antigens of invasive *Klebsiella pneumoniae* in a multi-country collection. *Front. Microbiol.* **11**, 1249 (2020).
- Yang, Q. et al. Emergence of ST11-K47 and ST11-K64 hypervirulent carbapenem-resistant *Klebsiella pneumoniae* in bacterial liver abscesses from China: a molecular, biological, and epidemiological study. *Emerg. Microbes Infect.* **9**(1), 320–331 (2020).
- Horváth, M. et al. Identification of a newly isolated lytic bacteriophage against K24 capsular type, carbapenem resistant *Klebsiella pneumoniae* isolates. *Sci Rep.* **10**(1), 5891 (2020).
- Shamanna, V. et al. Geographical distribution, disease association and diversity of *Klebsiella pneumoniae* K/L and O antigens in India: roadmap for vaccine development. *Microbial Genomics.* **10**(7), 001271 (2024).
- Spadar, A., Perdigão, J., Campino, S. & Clark, T. G. Large-scale genomic analysis of global *Klebsiella pneumoniae* plasmids reveals multiple simultaneous clusters of carbapenem-resistant hypervirulent strains. *Genome Med.* **15**(1), 3. <https://doi.org/10.1186/s13073-023-01153-y> (2023).
- Di Pilato, V., Pollini, S., Miriagou, V., Rossolini, G. M. & D'Andrea, M. M. Carbapenem-resistant *Klebsiella pneumoniae*: the role of plasmids in emergence, dissemination, and evolution of a major clinical challenge. *Expert Rev. Anti Infect. Ther.* **22**(1–3), 25–43 (2024).
- Lan, P., Jiang, Y., Zhou, J. & Yu, Y. A global perspective on the convergence of hypervirulence and carbapenem resistance in *Klebsiella pneumoniae*. *J. Global Antimicrob. Resist.* **25**, 26–34 (2021).
- Tian, D. et al. Prevalence of hypervirulent and carbapenem-resistant *Klebsiella pneumoniae* under divergent evolutionary patterns. *Emerg. Microbes Infect.* **11**(1), 1936–1949 (2022).
- Kain, M. J. et al. The rapid emergence of hypervirulent *Klebsiella* species and *Burkholderia pseudomallei* as major health threats in Southeast Asia: The urgent need for recognition as neglected tropical diseases. *Trop. Med. Infect. Dis.* **9**(4), 80 (2024).
- Elias, R. et al. Emergence of KPC-3- and OXA-181-producing ST13 and ST17 *Klebsiella pneumoniae* in Portugal: genomic insights on national and international dissemination. *J Antimicrob Chemother.* **78**(5), 1300–1308 (2023).
- Spadar, A. et al. Genomic epidemiological analysis of *Klebsiella pneumoniae* from Portuguese hospitals reveals insights into circulating antimicrobial resistance. *Sci Rep.* **12**(1), 13791 (2022).
- Aslam, B. et al. Distribution and genetic diversity of multi-drug-resistant *Klebsiella pneumoniae* at the human–animal–environment interface in Pakistan. *Front Microbiol.* **13**, 898248 (2022).
- Guo, M.-Q. et al. Genomic epidemiology of hypervirulent carbapenem-resistant *Klebsiella pneumoniae* at Jinshan local hospital, Shanghai, during 2014–2018. *J. Microbiol. Immunol. Infect.* **57**(1), 128–137 (2024).
- Andrews, S. FastQC: a quality control tool for high throughput sequence data. *F1000Res* **7**, 1338 (2018).
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**(15), 2114–2120 (2014).
- Wood, D. E. & Lu, J. Langmead BJB Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**(1), 257 (2019).
- Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* **19**(5), 455–477 (2012).
- Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. J. B. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**(8), 1072–1075 (2013).
- Bortolaia, V. et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* **75**(12), 3491–3500 (2020).
- Carattoli A, Hasman H, protocols. PlasmidFinder and in silico pMLST: identification and typing of plasmid replicons in whole-genome sequencing (WGS). *Methods Mol Biol.* 2020:285–94.
- Chen L, Zheng D, Liu B, Yang J, Jin QJNar. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res.* 2016;44(D1):D694–D7.
- Lam, M. M. et al. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nature Comm.* **12**(1), 4188 (2021).
- Seemann, T. J. B. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**(14), 2068–2069 (2014).
- Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* **52**(W1), W78–W82 (2024).
- Alam, M. et al. Tackling antimicrobial resistance in primary care facilities across Pakistan: current challenges and implications for the future. *J. Infect. Public Health* **16**, 97–110 (2023).

37. Gurung S, Kafle S, Dhungel B, Adhikari N, Thapa Shrestha U, Adhikari B, et al. Detection of OXA-48 gene in carbapenem-resistant *Escherichia coli* and *Klebsiella pneumoniae* from urine samples. *Infection and Drug Resistance*. 2020;2311–21.
38. RahmatUllah, S. et al. Exploring the resistome, virulome, and mobilome of multidrug-resistant *Klebsiella pneumoniae* isolates: deciphering the molecular basis of carbapenem resistance. *BMC Genomics* 25(1), 408 (2024).
39. Qamar MU, Sierra R, Jabeen K, Rizwan M, Rashid A, Dar YF, et al. Genomic characterization of plasmids harboring bla NDM-1, bla NDM-5, and bla NDM-7 carbapenemase alleles in clinical *Klebsiella pneumoniae* in Pakistan. *Microbiology Spectrum*. 2025:e02359–24.
40. Enany, S., Zakeer, S., Diab, A. A., Bakry, U. & Sayed, A. A. Whole genome sequencing of *Klebsiella pneumoniae* clinical isolates sequence type 627 isolated from Egyptian patients. *PLoS ONE* 17(3), e0265884 (2022).
41. Peirano, G., Chen, L., Kreiswirth, B. N. & Pitout, J. D. Emerging antimicrobial-resistant high-risk *Klebsiella pneumoniae* clones ST307 and ST147. *Antimicrob. Agents Chemotherapy*. <https://doi.org/10.1128/aac.01148-20> (2020).
42. Di Pilato, V. et al. Resistome and virulome accretion in an NDM-1-producing ST147 sublineage of *Klebsiella pneumoniae* associated with an outbreak in Tuscany, Italy: a genotypic and phenotypic characterisation. *The Lancet Microbe*. 3(3), e224–e234 (2022).
43. Lee, C.-R. et al. Antimicrobial resistance of hypervirulent *Klebsiella pneumoniae*: epidemiology, hypervirulence-associated determinants, and resistance mechanisms. *Front Cell Infect Microbiol*. 7, 483 (2017).
44. Luo, K. et al. Nosocomial infection by *Klebsiella pneumoniae* among neonates: a molecular epidemiological study. *J Hosp Infect*. 108, 174–180 (2021).
45. Tian, X. et al. First description of antimicrobial resistance in carbapenem-susceptible *Klebsiella pneumoniae* after imipenem treatment, driven by outer membrane remodeling. *BMC Microbiol*. 20, 1–11 (2020).
46. Perdigão, J. et al. Genomic epidemiology of carbapenemase producing *Klebsiella pneumoniae* strains at a northern Portuguese hospital enables the detection of a misidentified *Klebsiella variicola* KPC-3 producing strain. *Microorganisms*. 8(12), 1986 (2020).
47. Legese, M. H. et al. Molecular epidemiology of extended-spectrum beta-lactamase and AmpC producing Enterobacteriaceae among sepsis patients in Ethiopia: a prospective multicenter study. *Antibiotics (Basel)*. 11(2), 131 (2022).
48. Imtiaz, W., Dasti, J. I. & Andrews, S. C. Draft genome sequence of a carbapenemase-producing (NDM-1) and multidrug-resistant, hypervirulent *Klebsiella pneumoniae* ST11 isolate from Pakistan, with a non-hypermucoviscous phenotype associated with rmpA2 mutation. *J Glob Antimicrob Resist*. 25, 359–362 (2021).
49. Scavuzzi, A. M. L. et al. Occurrence of qnrB1 and qnrB12 genes, mutation in gyrA and ramR, and expression of efflux pumps in isolates of *Klebsiella pneumoniae* carriers of bla KPC-2. *J Med Microbiol*. 66(4), 477–484 (2017).
50. Kareem SM, Al-Kadmy IM, Kazaal SS, Mohammed Ali AN, Aziz SN, Makharita RR, et al. Detection of gyrA and parC mutations and prevalence of plasmid-mediated quinolone resistance genes in *Klebsiella pneumoniae*. *Infect Drug Resist*. 2021:555–63.
51. Zarras, C. et al. Genetic characterization of carbapenem-resistant *Klebsiella pneumoniae* clinical isolates in a tertiary hospital in Greece, 2018–2022. *Antibiotics (Basel)*. 12(6), 976 (2023).
52. Turton, J. F., Perry, C., McGowan, K., Turton, J. A. & Hope, R. J. *Klebsiella pneumoniae* sequence type 147: a high-risk clone increasingly associated with plasmids carrying both resistance and virulence elements. *J Med Microbiol*. 73(4), 001823 (2024).
53. Moses VK, Kandi V, Bharadwaj VG, Suvvari TK, Podaralla EJC. Molecular Characterization of *Klebsiella pneumoniae* Clinical Isolates Through Whole-Genome Sequencing: A Comprehensive Analysis of Serotypes, Sequence Types, and Antimicrobial and Virulence Genes. *Cureus*. 2024;16(4).
54. Fraise, A. J. Biocide abuse and antimicrobial resistance—a cause for concern?. *J Antimicrob Chemother*. 49(1), 11–12 (2002).
55. Mbelle, N. M. et al. Pathogenomics and evolutionary epidemiology of multi-drug resistant clinical *Klebsiella pneumoniae* isolated from Pretoria, South Africa. *Sci Rep*. 10(1), 1232 (2020).
56. Wyres, K. L., Lam, M. M. & Holt, K. E. Population genomics of *Klebsiella pneumoniae*. *Nat Rev Microbiol*. 18(6), 344–359 (2020).
57. Navon-Venezia, S., Kondratyeva, K. & Carattoli, A. *Klebsiella pneumoniae*: a major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiol Rev*. 41(3), 252–275 (2017).
58. Attalla, E. T., Khalil, A. M., Zakaria, A. S., Baker, D. J. & Mohamed, N. M. Genomic characterization of colistin-resistant *Klebsiella pneumoniae* isolated from intensive care unit patients in Egypt. *Ann Clin Microbiol Antimicrob*. 22(1), 82 (2023).
59. Zhao, L. et al. Molecular epidemiology of antimicrobial resistance, virulence and capsular serotypes of carbapenemase-carrying *Klebsiella pneumoniae* in China. *Antibiotics (Basel)*. 11(8), 1100 (2022).
60. Turton, J. F., Payne, Z., Micah, K. & Turton, J. A. Capsular type K54, clonal group 29 and virulence plasmids: an analysis of K54 and non-K54 closely related isolates of *Klebsiella pneumoniae*. *Epidemiol Infect*. 146(14), 1813–1823 (2018).

Acknowledgements

We thank the Office of Research, Innovation & Commercialization (ORIC), Khyber Medical University, Peshawar, Pakistan, for financial support in providing reagents and other research-related materials.

Author contributions

A.; I.U Conceptualization, Methodology: Investigation, writing—original draft, W.H.; J.E.P.; I.U.H; Reviewing, Editing and Analysis, O.C.M.; Z.U.D; Visualization and Analysis, A.N.; M.Z.P.; G.F.R.; Software, Investigation, Analysis, S.C.; S.A.; Resources and validation, T.A.K.; T.G.C.; supervision, project administration, and funding. All authors have read and agreed to the published version of the manuscript.

Funding

TGC and SC are funded by the UKRI (BBSRC BB/X018156/1; MRC MR/X005895/1; EPSRC EP/Y018842/1).

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval

This was granted by the Institutional Ethics Committee of the Institute of Pathology and Diagnostic Medicine on 1 June 2022 (Ref. No: KMU/IPDM/IEC/202229). All experimental protocols were approved, and all methods were performed in accordance with the relevant guidelines and regulations.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-27122-6>.

Correspondence and requests for materials should be addressed to T.A.K. or T.G.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025